MOTION IMAGERY STANDARDS BOARD	MISB ST 1010.1
STANDARD	
Generalized Standard Deviation and Correlation Coefficient Metadata	**27 February 2014**

1 Scope

This Standard (ST) defines a bit-efficient Standard Deviation and Correlation Coefficient data method. This ST is dependent on further context from an invoking Standard.

The method utilizes the symmetry of the variance-covariance matrix and the fixed data range of the correlation coefficients for "compression;" therefore, this method cannot be used in more generic matrix cases. This ST also defines a method for transmitting sparse standard deviation and correlation coefficient data along with standard deviation-only data. Finally, this document defines standard notation and usage of standard deviation and correlation coefficient information for an invoking Standard or method.

2 References

2.1 Normative Reference

The following references and the references contained therein are normative.

[1] MISB ST 1201.1, Floating Point to Integer Mapping, Feb 2014
[2] MISB ST 0107.2, Bit and Byte Order for Metadata in Motion Imagery Files and Streams, Feb 2014

3 Revision History

Revision	Date	Summary of Changes
ST 1010.1	2/27/2014	• Promoted to Standard

4 Abbreviations and Acronyms

2D Two Dimensional
ECEF Earth Centered Earth Fixed
FLP Floating Length Pack
KLV Key-Length-Value
LS Local Set
US Universal Set

5 Introduction

A variance-covariance matrix is composed of a number of standard deviations and correlation coefficients. The metadata required to describe this variance-covariance matrix is needed for proper photogrammetric processing, and because photogrammetric processing could occur with data generated at the motion imagery frame rate, it is absolutely critical the standard deviation and correlation coefficient data are transmitted with as few bytes as possible. The techniques described below are a blend of bit efficiency and flexibility for multiple uses of transmitting a variance-covariance matrix to end-user applications. An example of a variance-covariance matrix is given below in Equation 1.

$$Q = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix}$$ **Equation 1**

This matrix may be divided into two types of parameters: (1) standard deviations; and (2) correlation coefficients. These two parameters are in Equation 2, where the standard deviations are represented with σ and the correlation coefficients are represented with ρ.

$$Q = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix} \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{21} & 1 & \rho_{23} \\ \rho_{31} & \rho_{32} & 1 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix}$$ **Equation 2**

Correlation coefficients are mathematical values that describe how two random variables behave relative to one another. They are derived from a variance-covariance matrix produced from a stochastic process, such as least-squares adjustments or Kalman filtering. The correlation coefficient values are bounded between negative-one and positive-one [-1.0, 1.0]. Typical variance-covariance matrices are symmetrical about the main diagonal; for example, the value with subscripts (1, 2) is identically equal to the value with subscripts (2, 1). This symmetry reduces the number of elements required to reconstruct the full variance-covariance matrix. The relationship between correlation coefficients, standard deviation, and covariance may be computed (when corresponding values are not zero) per Equation 3.

$$\rho_{ij}\sigma_i\sigma_j = \sigma_{ij}$$ **Equation 3**

This ST defines an efficient method to represent the standard deviations and correlation coefficients in a Floating Length Pack (FLP), which may be used in KLV groupings like Universal Sets, Local Sets, etc.

6 Standard Deviation and Correlation Coefficient Metadata

This section provides the background, describes the development, and formalizes the bit efficient packaging of standard deviation and correlation coefficient data.

6.1 Overview and Definitions

A two dimensional (2D) standard deviation and correlation coefficient matrix is “encoded” into a linear block of data. A receiver typically decodes this data block back to a 2D standard deviation and correlation coefficient matrix to support further uncertainty propagation. Note the standard deviation and correlation coefficient matrix must be converted to a variance-covariance matrix to perform any further uncertainty propagation. The matrix of standard deviations and correlation coefficients is a representation of this matrix that helps simplify the description of the transmission process. Figure 1 illustrates the overall flow and aspects of the encoding. The Source Matrix of Standard Deviations and Correlation Coefficients Properties, Packaging and Compression, Encoding/Decoding, and the Reconstructed Matrix of Standard Deviations and Correlation Coefficients are discussed next.

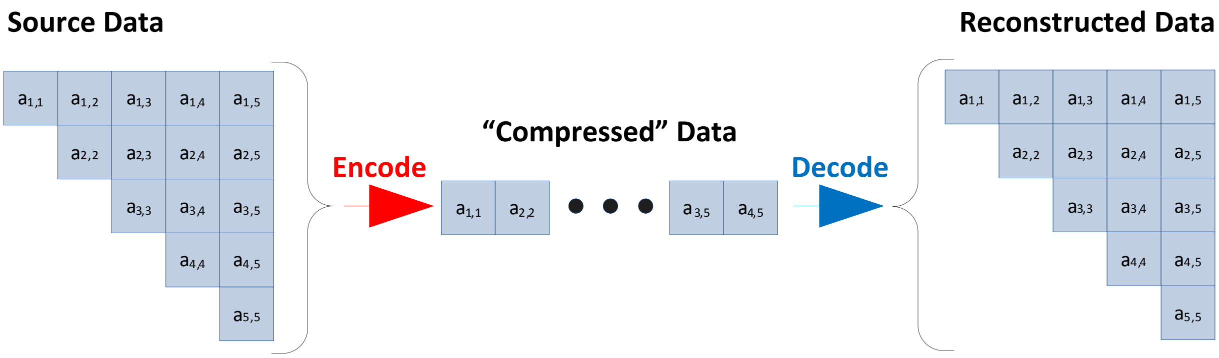

Figure 1: Encode-to-Decode Process

6.2 Standard Deviation and Correlation Coefficient Matrix Properties

The standard deviation and correlation coefficient matrix is a square symmetric matrix, where the diagonal values of the matrix contain the standard deviation values and the off-diagonal values contain the correlation coefficient values. Because of the symmetry in the correlation coefficients about the diagonal, the lower triangular values are omitted as shown in Figure 2.

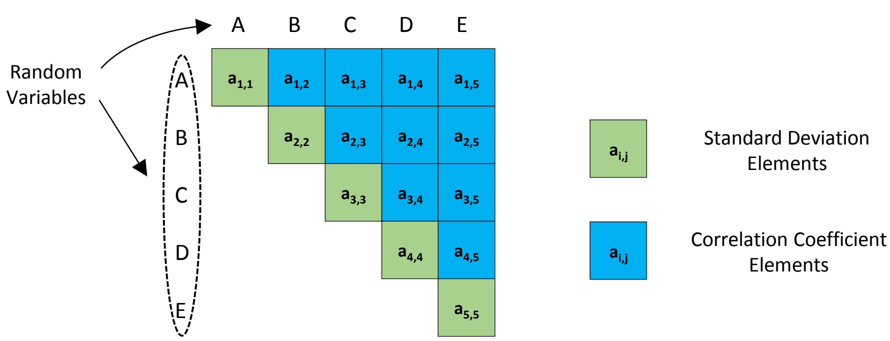

Figure 2: Standard Deviation and Correlation Coefficient Matrix

Each row of the standard deviation and correlation coefficient matrix in Figure 2 represents a random variable compared to a list of other random variables, which are represented by each column. A random variable is created when a measurement or computation is made about a specific phenomenon. For example, a set of measurements for the ECEF X, Y and Z position of a sensor. The uncertainty propagation parameters for these three items in a 3x3 matrix contain two pieces of information about those measurements: (1) the standard deviation of each individual element (how well the measurement was made – i.e. the error); and (2) the correlation coefficient or relationship of one measurement to another (i.e. the uncertainty in the X measurement is impacted by the uncertainty in the Y measurement).

The matrix in Figure 2 shows the upper triangular portion of the standard deviation and correlation coefficient matrix created by a collection of measurements. However, there are cases when (a) individual measurements may not have been made, or (b) the correlation coefficient between two values is unknown (or zero). When one or more measurements are not made, those rows (and corresponding columns) are set to zero, which is the equivalent of removing those rows (and columns) from the overall matrix. This in turn reduces the data sent to the receiver. When the correlation coefficient value between two variables is unknown or zero, that value is eliminated from the transmission, which also saves bandwidth causing the matrix to be sparse. Figure 3 illustrates these two cases of a row/column elimination and sparse matrix.

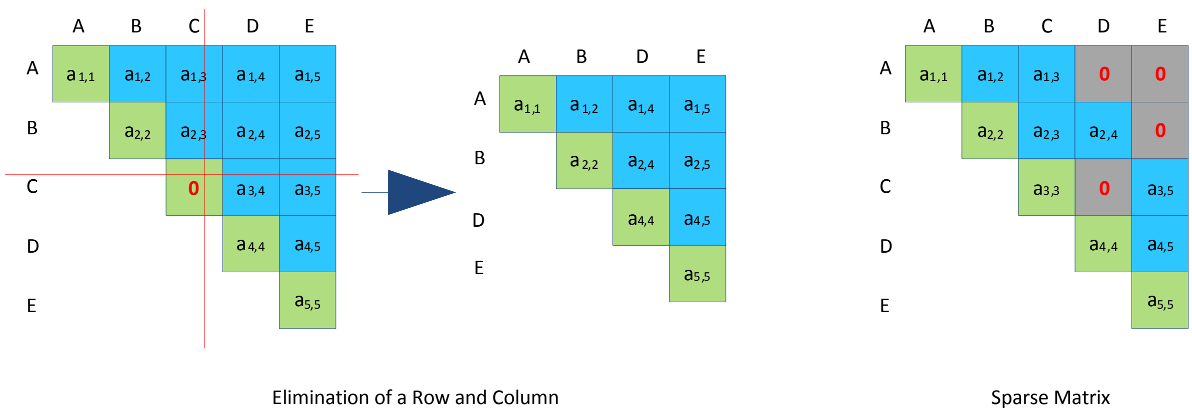

Figure 3: Reduced and Sparse Standard Deviation and Correlation Coefficient Matrix

6.3 Packaging and Compression

6.3.1 Symmetry

One property of a matrix of standard deviations and correlation coefficients is symmetry about the diagonal, which means all values in the upper triangular area are equal to all of those in the lower triangular area. By eliminating the lower triangular portion of the matrix, there is a major reduction in the bytes for representation; this increases bit efficiency. This ST takes advantage of the symmetry by only transmitting the upper triangular area, where the lower triangular values are constructed as needed. Figure 4 indicates the data encoded into KLV.

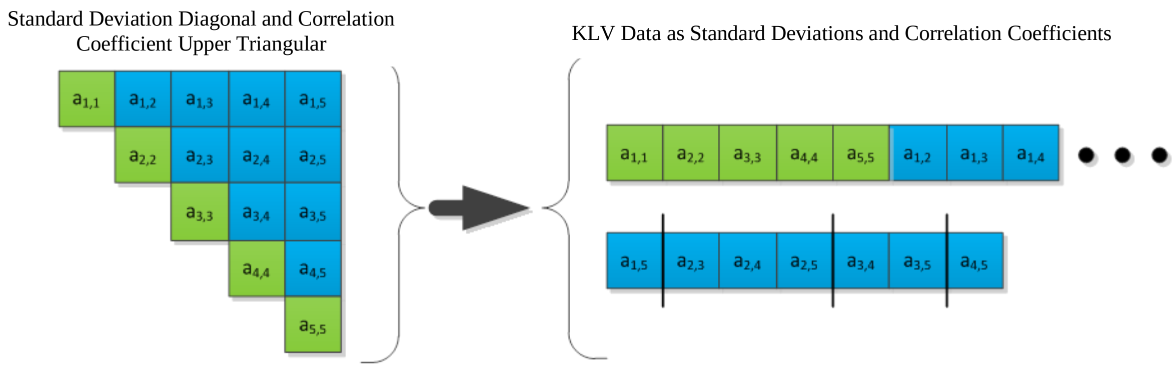

Figure 4: Standard Deviation and Correlation Coefficient Data Prepared for KLV Encoding

6.3.2 Type and Size

The standard deviation and correlation coefficient data are formatted for transmission differently; each will have their own size and type (Float, Integer, etc.).

Requirement	
ST 1010.1-01	When the Standard or method invoking MISB ST 1010 describes a different byte size or data type for the standard deviation metadata parameters, the greatest or most conservative value shall be used for encoding to preserve precision.
ST 1010.1-02	When the Standard or method invoking MISB ST 1010 describes a different byte size for the correlation metadata parameters, the greatest or most conservative value shall be used for encoding to preserve precision.

As an example, consider an invoking Standard or method that specifies ECEF X, Y, and Z positions for transmission, and standard deviations are required to describe the uncertainty of these values. The invoking Standard or method indicates the variable type and size. Assume that some of the standard deviations are represented as four-byte Floating Point numbers, and others as two-byte ST 1201[1] mapped integers. As all standard deviation values in the FLP must have the same length, the two-byte integer values are coded as four-byte Floating Point numbers; this preserves the precision of all the values (see Figure 5).

Likewise, the correlation coefficient values must all have the same number of bytes for each value within a FLP. However, because the range for a correlation coefficient value is always [-1.0, 1.0], the correlation coefficient values are mapped into integers according to ST 1201. Continuing with the example, assume all the correlation coefficients for the ECEF X, Y, and Z positions are represented using two-byte values. The completed stream of data values is as shown in Figure 5.

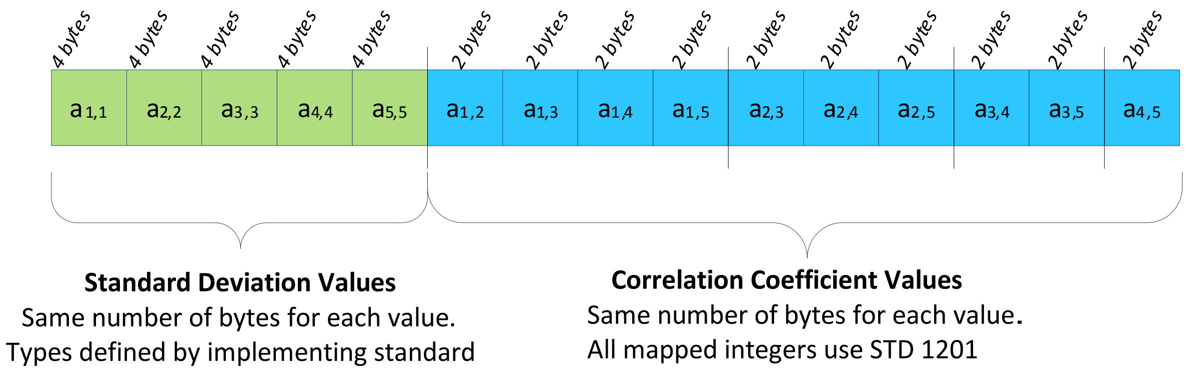

Figure 5: Standard Deviation and Correlation Coefficient Bytes

In order to interpret this data and parse the correct number of bytes, the data size for the standard deviations (S_{len}) and correlation coefficients (C_{len}) are included in the data stream; these values precede the data values. The data sizes for the standard deviations and the correlation coefficient range from 0 to 6 bytes, where zero is a special case discussed below. The data sizes are encoded into a single byte with each size (S_{len} and C_{len}) using three bits; the remaining bit (C_s) is used for another purpose described in Section 6.3.4. Figure 6 shows the bit layout for this byte; bits 0-2 are used to define the C_{len} value, bits 4-6 are used to define the S_{len} value; bits 3 is used as a flag which is discussed in Section 6.3.4.

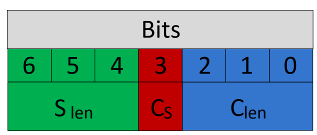

Figure 6: Standard Deviation and Correlation Coefficient Bit Assignment

The three-bit values provide for a maximum size of seven bytes. A data size with a value of zero means there are no values of that type (standard deviation or correlation coefficient) of data transmitted. For example, if the correlation coefficient data length value is zero, then only standard deviation data is transmitted. The examples in Figure 7 illustrate the addition of the length information byte; the top graphic shows both standard deviation (4 bytes per value) and correlation coefficient (2 bytes per value) information, and the bottom graphic shows only standard deviation (4 bytes per value) information (correlation coefficient of 0 bytes per value).

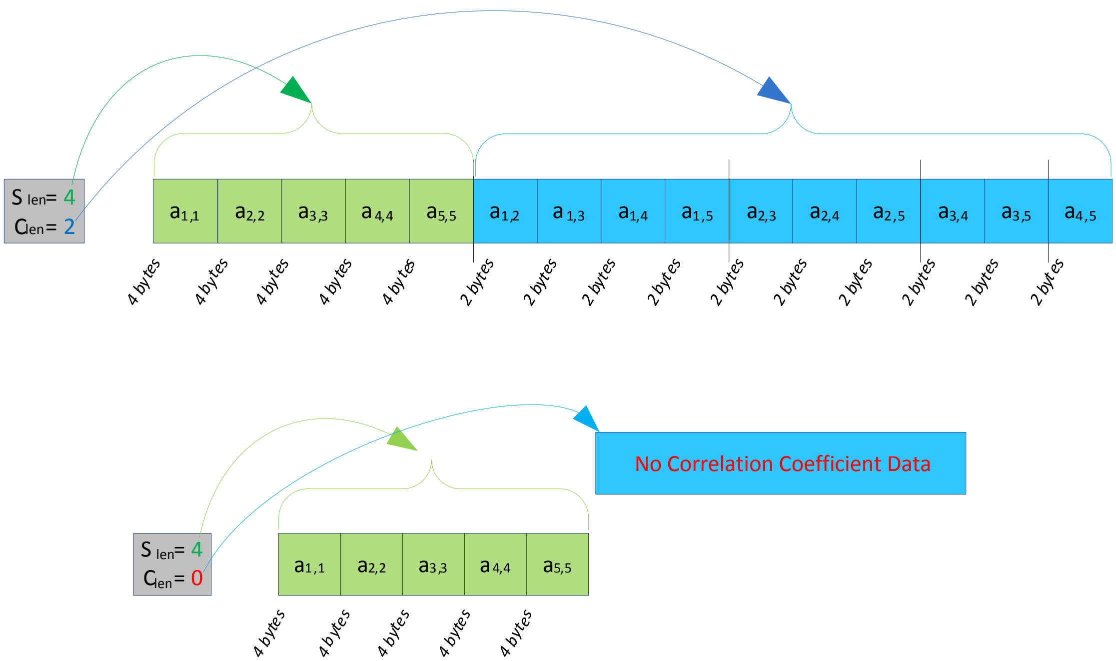

Figure 7: Standard Deviation and Correlation Coefficient KLV Data with Bit Information

6.3.3 Matrix Dimension

In order to reconstruct a matrix, the dimension of the matrix is required before parsing can begin. The matrix dimension is included in the pack as a BER OID encoded value. Therefore, there is no theoretical limit to the dimension size of the matrix. Figure 8 shows how the matrix dimension N is used to define the lengths of each row.

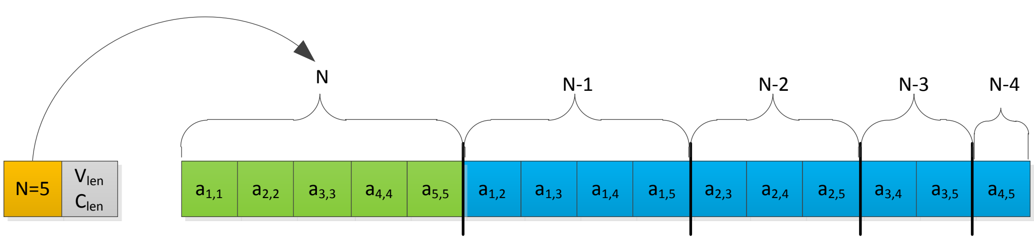

Figure 8: Standard Deviation and Correlation Coefficient KLV Data with Bit Information and Matrix Dimension

6.3.4 Bit Vector (Sparse Representation)

The representation in Figure 8 shows how to build a full matrix of standard deviations and correlation coefficients once given a dimension N; however, there are cases where either rows (and corresponding columns) are eliminated, or correlation coefficient values are unknown or zero (i.e. sparse) – see Section 6.5 for more information. To provide for sparse matrices, a bit vector value is optionally included, where each bit represents whether a value is transmitted or not. The inclusion of the bit vector is controlled by the C_S bit, which is part of the S_{len} and C_{len} byte construction (see Figure 6).

Requirement	
ST 1010.1-03	A full matrix of Standard Deviation values shall always be present in the Standard Deviation and Correlation Coefficient Floating Length.

Table 1: Bit Vector Status based on Length of Correlation Coefficients

C_{len}	Correlation Coefficient Values	C_s	Bit Vector
0	Not present in the data	X	Not Needed
> 0	Full Set Present	0	Not Needed
> 0	Sparse set present in the data	1	Indicates Values Present

The bit vector is only included if there is sparse data for the correlation coefficients. When present it is the same order and length as the correlation coefficient information. The byte size of the bit vector is computed by having one bit for each potential correlation coefficient value, rounded to the nearest byte (see top graphic in Figure 9). If any correlation coefficient data is absent from the pack, the bit vector only applies to the data that is present; the bit vector length is adjusted depending on what data is included. Equation 4 is used to compute the length of the bit vector.

$$V_{bytes} = \left\lceil \frac{C_S\left(\frac{N(N-1)}{2}\right)}{8} \right\rceil = \left\lceil \frac{C_S N(N-1)}{16} \right\rceil \qquad \textbf{Equation 4}$$

Where: $\lceil y \rceil$=ceiling(y) (i.e. round up)

As an example, Figure 9 shows a full matrix with the values to be sent along with several correlation coefficient values not populated. Values sent are indicated with a “1” and values not sent a “0”. The collection of ones and zeros comprise the bit vector. The calculation shows two bytes are needed to convey the bit vector. The bottom graphic shows the resulting block of data that would be transmitted.

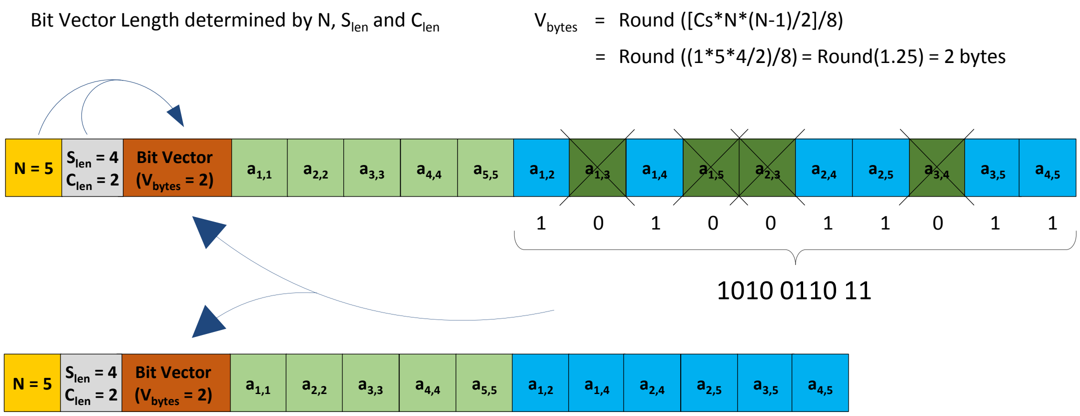

Figure 9: Standard Deviation and Correlation Coefficient KLV Data with Bit Information, Dimension Size, and Bit Mask

If only non-sparse standard deviation data is transmitted without any of the correlation coefficient values, then C_{len} is set to zero, C_S is set to zero, and the bit vector is not included at all; this is illustrated in Figure 10.

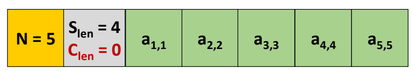

Figure 10: Standard Deviation Only KLV Data without Bit Mask

These examples demonstrate the flexibility and the efficiency of this method.

6.4 Standard Deviation and Correlation Coefficient Metadata FLP

The floating length pack (FLP) of standard deviation and correlation coefficient information is composed of individual items combined into one block of binary data. Five elements, in the order specified in Table 2, comprise the FLP: Matrix Size, Parse Control Byte, Bit Vector, Standard Deviation Elements, and the Correlation Coefficient Elements.

The invoking Standard or method specifies the Matrix Size N that corresponds to the N parameters in the invoking standard that have uncertainty information. These N parameters are identified by their KLV Tag in the invoking standard.

The Standard Deviation and Correlation Coefficient FLP affords efficiencies when a matrix is sparse, or lacking correlation values; attention to how the five elements in the FLP are populated can reduce the data required to transmit standard deviation and correlation coefficient information. For instance, by reordering the data set parameters in a KLV stream elements may be removed, which effectively sets a row and column for a value that does not contain a standard deviation to zero, thereby, reducing bytes. The Parse Control Byte and Bit Vector is set by the invoking Standard or method to characterize and define the contents of data present in the FLP.

Multiple Standard Deviation and Correlation Coefficient FLPs are allowed per stream. Each FLP describes a number of N parameters in the invoking Standard or method. Option 3 in the example section 6.5.1.3 illustrates this use case.

Table 2: Standard Deviation and Correlation Coefficient FLP Components Overview

FLP Key		Name	
06 0E 2B 34 02 05 01 01 0E 01 03 03 21 00 00 00 (CRC 64882)		**Standard Deviation and Correlation Coefficient FLP**	
Name	**Type**	**Size**	**Description**
Matrix Size	BER OID Integer	Variable	Describes the dimension, N, of the matrix of standard deviations and correlation coefficients that is being transmitted. This value is a BER OID encoded count of the number of rows or columns in the square matrix.
Parse Control Byte	Packed Value	1 Byte	Contains three values (S_{len}, C_S, C_{len}) encoded into a single byte. C_S is a single bit that indicates if the correlation coefficient values are sparsely represented or not; if the C_S bit is set to true (1) then the Bit Vector (next element) is included in this data block. The S_{len} and C_{len} values indicate the number of bytes used for the standard deviation values and the correlation coefficient values, respectively. If S_{len} is zero, then there are no standard deviation values in the data block. If C_{len} is zero, then there are no correlation coefficient values in the data block. See Figure 6 for how the bits are assigned: S_{len}= bits 4-6, C_S = bits 3, C_{len} = bits 0-2.
Bit Vector	Array of Bits	Variable	An optional list of bits that denote which values of the correlation coefficients are being transmitted in this package. The size of this list is computed using Equation 4. This bit vector along with the Matrix Size is used to determine where values populate the two dimensional standard deviation and correlation coefficient array. The existence of this data is controlled by the values contained in C_S and C_{len}.
Standard Deviation Elements	Array of Values	Variable	A list of values where each value has a byte size specified by S_{len}. The number of values is depended on the Matrix Size, specified by the invoking Standard or method, and the bits in the Bit Vector.
Correlation Coefficient Elements	Array of Integer Mapped Values	Variable	A list of values where each value has a byte size specified by C_{len}. The number of values is depended on the Matrix Size and the bits in the Bit Vector. Each value is a floating point-to-integer mapped value using MISB ST 1201[1] for mapping the values. The range of each value is [-1.0, 1.0], so the mapping function, as specified by ST 1201, is IMAPB(-1.0, 1.0, C_{len}).

6.5 Invoking ST 1010

The Standard Deviation and Correlation Coefficient FLP is a construct invoked by another Standard or method. The following are requirements for the use of the Standard Deviation and Correlation Coefficient FLP.

Requirement	
ST 1010.1-04	The Standard Deviation and Correlation Coefficient FLP shall contain all of the elements in MISB ST 1010 Table 2 Standard Deviation and Correlation Coefficient FLP.

ST 1010.1-05	When a MISB KLV parameter requires standard deviation and/or correlation coefficient information, the invoking Standard or method that specifies that KLV parameter shall populate the five elements described in the Standard Deviation and Correlation Coefficient FLP in accordance with MISB ST 1010 Table 2.
ST 1010.1-06	The ordering of the parameters in a standard deviation and correlation coefficient matrix shall correspond to the same ordering of the invoking Standard or method KLV LS Tag numbers that correspond to the N parameters.
ST 1010.1-07	The Standard Deviation and Correlation Coefficient FLP that describes a specific set of parameters in the invoking Standard or method shall follow those specific parameters in the invoking Standard or method in the KLV stream.
ST 1010.1-08	All metadata shall be expressed in accordance with MISB ST 0107[2].

6.5.1 Invoking Example - Informative

Assume a Standard or method XXYY defines a Local Set of 12 elements, where only 6 of the 12 require standard deviations and correlation coefficients. There are three options for invoking ST 1010: (1) placing the Standard Deviation and Correlation Coefficient FLP at the end of the parent local set without considering order, (2) ordering the parameters into random variables and constants, and (3) further ordering of the random variables to minimize total data (i.e. multiple instances of the Standard Deviation and Correlation Coefficient FLP). These three options are arranged from least efficient to most efficient (theoretically).

6.5.1.1 Option 1

Option 1 requires no rearranging of the tag IDs present in the data set's KLV stream. At the end of the twelfth tag number, the Standard Deviation and Correlation Coefficient FLP is inserted into the KLV stream (see Table 3).

Table 3: Option 1 – Example Local Set

Tag ID Number	Name	Standard Deviation and Correlation Coefficient Applicable (For Example Purposes)
1	Sensor ECEF Position Component X	YES
2	Sensor ECEF Position Component Y	YES
3	Sensor ECEF Position Component Z	YES
4	Sensor ECEF Velocity Component X	NO
5	Sensor ECEF Velocity Component Y	NO
6	Sensor ECEF Velocity Component Z	NO
7	Sensor Heading Angle	YES
8	Sensor Pitch Angle	YES
9	Sensor Roll Angle	YES
10	Sensor Absolute Heading Rate	NO
11	Sensor Absolute Pitch Rate	NO
12	Sensor Absolute Roll Rate	NO
ST 1010 Standard Deviation and Correlation Coefficient FLP N = 12		

6.5.1.2 Option 2

Option 2 rearranges the tag IDs, so the six items requiring standard deviations and correlation coefficients are sent together in the KLV stream (see Table 4).

Table 4: Option 2 – Example Local Set

Tag ID Number	Name	Standard Deviation and Correlation Coefficient Applicable (For Example Purposes)
1	Sensor ECEF Position Component X	YES
2	Sensor ECEF Position Component Y	YES
3	Sensor ECEF Position Component Z	YES
7	Sensor Heading Angle	YES
8	Sensor Pitch Angle	YES
9	Sensor Roll Angle	YES
ST 1010 Standard Deviation and Correlation Coefficient FLP N = 6		
4	Sensor ECEF Velocity Component X	NO
5	Sensor ECEF Velocity Component Y	NO
6	Sensor ECEF Velocity Component Z	NO
10	Sensor Absolute Heading Rate	NO
11	Sensor Absolute Pitch Rate	NO
12	Sensor Absolute Roll Rate	NO

6.5.1.3 Option 3

Option 3 has multiple Standard Deviation and Correlation Coefficient FLPs in the KLV stream. The 12 tag IDs may be strategically organized into multiple groups (see Table 5).

Table 5: Option 3 – Example Local Set

Tag ID Number	Name	Standard Deviation and Correlation Coefficient Applicable (For Example Purposes)
1	Sensor ECEF Position Component X	YES
2	Sensor ECEF Position Component Y	YES
3	Sensor ECEF Position Component Z	YES
ST 1010 Standard Deviation and Correlation Coefficient FLP N = 3		
4	Sensor ECEF Velocity Component X	NO
5	Sensor ECEF Velocity Component Y	NO
6	Sensor ECEF Velocity Component Z	NO
7	Sensor Heading Angle	YES
8	Sensor Pitch Angle	YES
9	Sensor Roll Angle	YES
ST 1010 Standard Deviation and Correlation Coefficient FLP N = 3		
10	Sensor Absolute Heading Rate	NO
11	Sensor Absolute Pitch Rate	NO
12	Sensor Absolute Roll Rate	NO